

Apparent limitations in the ability of CMIP5 climate models to simulate recent multi-decadal change in surface temperature: implications for global temperature projections

Scott Power¹ · François Delage¹ · Guomin Wang¹ · Ian Smith¹ · Greg Kociuba¹

Received: 22 December 2015 / Accepted: 18 August 2016 / Published online: 28 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Observed surface temperature trends over the period 1998–2012/2014 have attracted a great deal of interest because of an apparent slowdown in the rate of global warming, and contrasts between climate model simulations and observations of such trends. Many studies have addressed the statistical significance of these relatively short-trends, whether they indicate a possible bias in the model values and the implications for global warming generally. Here we re-examine these issues, but as they relate to changes over much longer-term changes. We find that on *multi-decadal* time scales there is little evidence for any change in the observed global warming rate, but some evidence for a recent temporary slowdown in the warming rate in the Pacific. This multi-decadal slowdown can be partly explained by a cool phase of the Interdecadal Pacific Oscillation and a short-term excess of La Niña events. We also analyse historical and projected changes in 38 CMIP climate models. All of the model simulations examined simulate multi-decadal warming in the Pacific over the past half-century that exceeds observed values. This difference cannot be fully explained by observed internal multi-decadal climate variability, even if allowance is made for an apparent tendency for models to underestimate internal multi-decadal variability in the Pacific. Models which simulate the greatest global warming over the past half-century also project warming that is among the highest of all models by the end of the twenty-first century, under both low

and high greenhouse gas emission scenarios. Given that the same models are poorest in representing observed multi-decadal temperature change, confidence in the highest projections is reduced.

Keywords Multi-decadal global temperature change · Pacific temperature · Interdecadal Pacific Oscillation · CMIP5 models · Climate sensitivity

1 Introduction

The behaviour of globally average surface temperatures over the period 1998 to 2012/2014 has attracted a great deal of attention, with questions being raised as to whether the observed values represented a serious hiatus (or pause) in global warming or whether they simply corresponded to a decadal-scale fluctuation associated with internal variability of the climate system (Lewandowsky et al. 2015, 2016). The Intergovernmental Panel on Climate Change (IPCC 2013; Flato et al. 2013) concluded that a reduction of the warming trend (1998–2012) was attributable, in roughly equal measure, to a cooling contribution from both internal climate variability, and external forcing comprising solar and volcanic activity. For example, Fyfe et al. (2016) estimated that the planet warmed at a rate of +0.170 °C per decade from 1972 to 2001, but at a slower rate of +0.113 °C per decade from 2000 to 2014, contrary to the expectations for an increase in global warming over time. However, an updated analysis by Karl et al. (2015) indicated no change in the global warming trend, while Lewandowsky et al. (2015) also found no substantive evidence for a “pause” or “hiatus”, maintaining that it was “statistically indistinguishable from previous fluctuations”.

Electronic supplementary material The online version of this article (doi:10.1007/s00382-016-3326-x) contains supplementary material, which is available to authorized users.

✉ Scott Power
s.power@bom.gov.au

¹ Bureau of Meteorology, Melbourne, VIC, Australia

Closely related to this issue is an apparent discrepancy between observed temperatures and those simulated by climate models. The Intergovernmental Panel on Climate Change (IPCC 2013; Flato et al. 2013) also noted that: “*There is very high confidence that [climate] models reproduce the general features of the global—scale annual mean surface temperature increase over the historical period, including the more rapid warming in the second half of the twentieth century*”. They also concluded, however, that “*most simulations of the historical period do not reproduce the observed reduction in global mean surface warming trend over the last 10–15 years*”. They concluded that these differences were due, to a substantial degree, to internal variability in the real world with possible contributions from forcing errors and, for some climate models, an overestimation of the response to increasing greenhouse gases (GHGs) and other anthropogenic forcings including sulphate aerosols (c.f. Bindoff et al. 2013). Their analysis also indicates most models tended to warm more than observations did from 1961–1990 to 1998–2012.

On the other hand, other studies (e.g. Brown et al. (2015); Marotzke and Forster (2015); England et al. (2015)) found no evidence for any systematic overestimate in the responses of the models, while Huber and Knutti (2014) concluded that “*... the reduced warming and mismatch between models and observations can, to a large extent, be explained by the combined effect of reduced forcing and natural variability, each of these components contributing about an equal amount*”.

It has been suggested that natural variability during this period comprised a cool phase of the Interdecadal Pacific Oscillation (IPO; or the closely Pacific Decadal Oscillation (PDO), Han et al. 2013). The IPO is a naturally occurring form of interdecadal climate variability centred in the Pacific Ocean (Power et al. 1999; Folland et al. 1999; Meehl and Arblaster 2012; Salinger et al. 2001; Henley et al. 2015; England et al. 2014; Kirtman et al. 2013). The IPO is known to have some effect on the variability of global temperatures on decadal time scales (Brown et al. 2014; Meehl et al. 2011, 2013). This variability is very largely absent from multi-model mean (MMM) values since any such cool and warm phases are uncorrelated across different model simulations such that their effects tend to cancel out. Consequently, several studies have focused on individual model simulations and their ability to match observed variability. Of particular interest here are the studies by Kosaka and Xie (2013), Watanabe et al. (2014), England et al. (2014), Meehl et al. (2014) and Kociuba and Power (2015), who focused on the ability of models to simulate Pacific Ocean variability. Kosaka and Xie (2013) showed that differences between modelled and observed decadal-scale trends could be largely removed in their climate model if the model was forced with observed sea surface temperatures

(SSTs) in the eastern tropical Pacific, rather than allowing the model to calculate its own SSTs. England et al. (2014) could largely simulate the observed changes by simply forcing their climate model with observed tropical Pacific winds. Watanabe et al. (2014) showed that approximately 30–50% of decadal-scale anomalies in global average temperatures could be simulated in their model when forced with observed tropical wind-stress anomalies. Kociuba and Power (2015) found that none of the 35 CMIP5 models they examined simulated the marked strengthening of the Walker circulation observed over the 33-year period 1980–2012, while Meehl et al. (2014) showed that some CMIP5 models could reproduce features of the hiatus period, given the correct initial conditions. They identified 10 out of 262 CMIP5 simulations that happened to match the slowdown in the observations as reflected in 15-year trends.

Other studies attempting to discriminate among the models and assess if there are any implications for projections for the end of the century include Huber and Knutti (2014), Risbey et al. (2014), England et al. (2015) and Marotzke and Forster (2015). Risbey et al. (2014) and England et al. (2015) argued that the recent discrepancies between models and observations were not that important, since there is little difference between the temperature projections for the end of the century from those models which can simulate features of the recent period and those that do not—the implication being that there should soon (i.e. post-2014) be a return to a warming trend similar to that predicted by the models (Huber and Knutti 2014).

With the benefit of hindsight, brought about by record warm temperatures associated with the 2015/2016 El Niño event, it is possible to better address some of the issues surrounding the so-called “hiatus” period. A great deal of attention has been given to trends over 15-year periods, especially 1998–2012, but here we are much more interested in the longer-term changes which better reflect anthropogenic climate change (Wehner and Easterling 2015; Lewandowsky et al. 2015). We consider the role of the Interdecadal Pacific Oscillation (IPO), the frequency of El Niño–Southern Oscillation (ENSO) events, and internal climate variability more broadly in accounting for the differences between observed and simulated multi-decadal changes in the Pacific and global temperatures. This is an important extension of previous work to longer time-scales. This means, for example, that the conclusions discussed above relating to e.g. 15-year trends may not necessarily apply. We will then consider what implications contrasts between model and observed multi-decadal temperature changes have for global and Pacific temperature projections for the coming century.

The data, models and methods used are described in Sect. 2. Results are presented in Sects. 3, 4, 5, 6. Section 3 provides an analysis of decadal, multi-decadal and

longer-term trends in global and Pacific temperatures using data up to 2016. Model means of the climate model simulations of the historical period are compared with observations in Sect. 4. The ability of internal, multi-decadal climate variability to account for the contrasts between the multi-decadal changes in the MMM and the observations is examined in Sect. 5. The implications the contrasts have for projections of global and Pacific temperature over the remainder of the twenty-first century are examined in Sect. 6. Results are summarised and discussed in Sect. 7.

2 Data and methods

Observed global annual average surface temperatures (1900–2015) are those generated by the Hadley Research Centre available from <http://www.metoffice.gov.uk/hadobs/hadcrut4/data/current/download.html> (HadCRUT4).

All temperatures are expressed as anomalies with respect to the 1961–1990 average (which is a standard reference period used by e.g. the World Meteorological Organization). We also frequently refer to the period 1998–2012, because this has been the focus of numerous studies dealing with the so-called hiatus, including the latest IPCC report (Flato et al. 2013).

At the time of writing, global values were available for much of 2016 and indicated that the end-of-year annual average was highly likely to exceed that of 2015 by about +0.2 °C (Schmidt, <https://mobile.twitter.com/ClimateOfGavin/status/731599988141248512>). Consequently, we estimate the 2016 value to be +0.95 °C.

Here we also refer to Pacific Ocean averages calculated for the box region 10 °S to 25 °N and 160 °E to 110 °W. This region has been chosen since it encompasses a node of the Interdecadal Pacific Oscillation, a pattern of SST variability at decadal time scales centred on the Pacific Ocean (England et al. 2014; Kirtman et al. 2013; Power et al. 1999; Folland et al. 1999; Meehl and Arblaster 2012; Salinger et al. 2001; Henley et al. 2015). It also represents a region where an apparent hiatus in warming trends is apparent.

A 2016 value for the Pacific box region (+0.71) was estimated based on the correlation with global average values and the fact that in previous years following major El Niño events, these values tend to decrease by about −0.4 °C. Note that the results are not affected by whether we restrict the analyses to 2015, when all available observations were available at the time of writing, or include the plausible estimates for 2016.

Eighty-three simulations from 38 models from the CMIP5 archive (Taylor et al. 2012) have been used to analyse the pre-industrial, the historical (1900–2005) and the future (2006–2100) periods under three (RCP2.6, RCP4.5 and RCP8.5) emissions scenarios. The historical runs

included forcing from increasing greenhouse gases, changing sulphate aerosols, major volcanic eruptions and solar variability (Moss et al. 2010; van Vuuren et al. 2011). The historical and future simulations were combined to produce model estimates for 1900–2100. Global means were calculated for each year using the same masks as the HadCRUT4 data set. For years after 2012, the 2012 masks were used. Note that the HadCRUT4 values may underestimate the actual rate of warming, because surface air temperatures over oceans are assumed to be the same as SSTs (Cowtan et al. 2015). The full list of models used is given in Supplementary Table 1.

IPO patterns are identified using the 2nd EOF of 13-year low pass filtered SSTs (Meehl and Arblaster 2012). We used gridded SST data (1877–2003) from the (MLOST) V3.5.324 data set (Vose et al. 2012) and model simulated values over the historical period 1900–2005. Data pre-processing was conducted as described by Jones et al. (2013). Both the observations and model values were re-gridded to a common $1.5^\circ \times 1.5^\circ$ grid. Very little Arctic data is included so that under sampling of the Arctic (Cowtan and Way 2014; Karl et al. 2015) does not affect the results.

While an IPO index can be defined as the EOF time series, it can also be more simply calculated as the linear combination of anomalies averaged over three distinct box regions (Henley et al. 2015). These index values, up to and including 2015, are available from <http://www.esrl.noaa.gov/psd/data/timeseries/IPOTPI>.

El Niño/La Niña events are defined according to NINO34 annual average temperatures. We firstly detrend the data (both observations and simulated values) by subtracting 31-year running averages from the raw values. El Niño/La Niña years are those in which the detrended anomalies exceed plus or minus one standard deviation (running 31-year values). An El-Niño-like index is either +1 for an El-Niño year, −1 for a La-Niña year and is zero otherwise (i.e. a ENSO neutral year). This standardization enables a comparison between different time series with different trends and different levels of variability and has been used successfully in previous studies (e.g. Power and Smith 2007).

3 Analysis of observations

Figure 1a shows observed global annual average surface temperature anomalies (relative to 1961–1990) for the period 1901 up to 2016. The years 1998–2012 are highlighted since this period has been the subject of so many studies. It can be seen that temperature anomalies could be described as relatively stagnant between 1997 and 2014. 31-year running averages indicate a fairly constant rate of

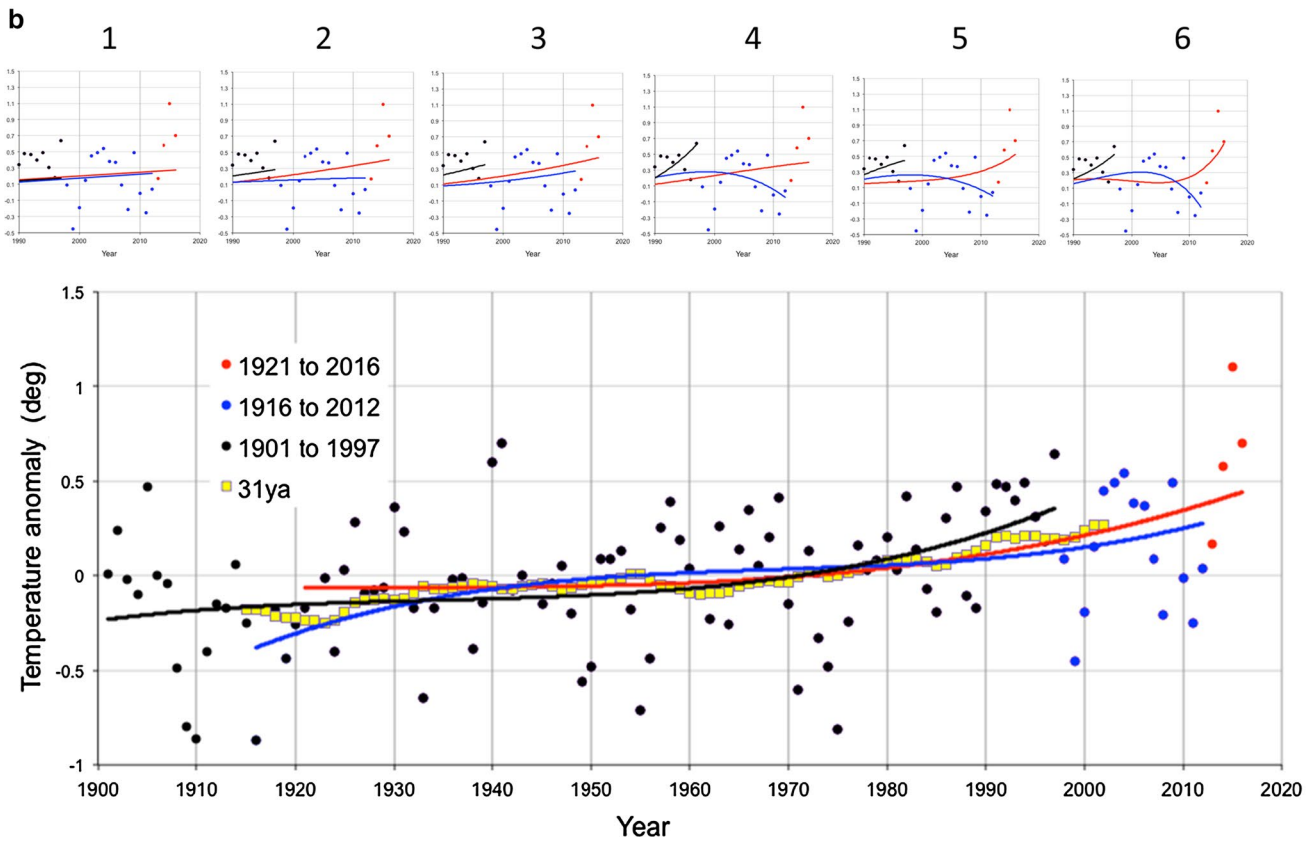
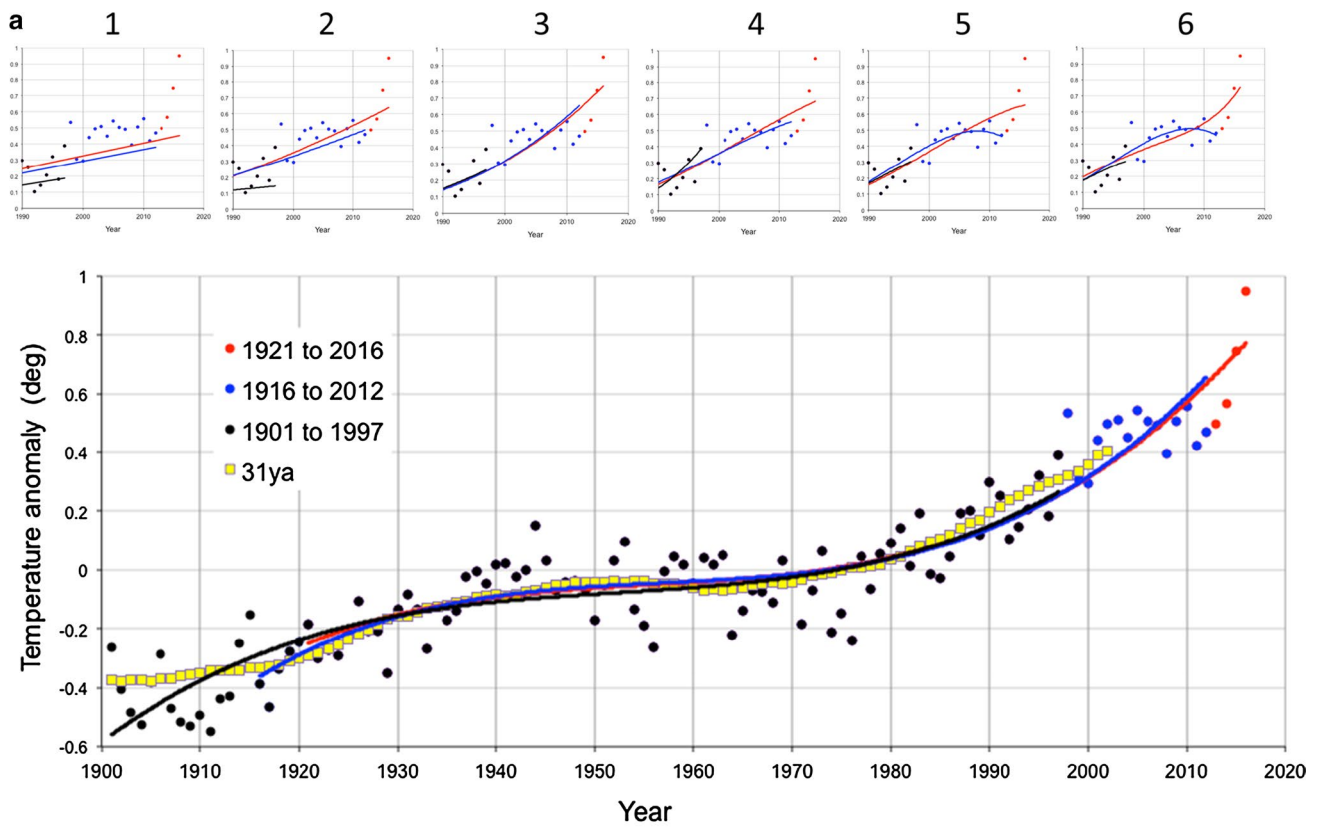


Fig. 1 Observed annual average surface temperatures 1921–2016 expressed as anomalies relative to the average for the period 1961–1990. The values over the period 1901–1997 are indicated by black circles, the values between 1998 and 2012 are indicated by blue circles, and the values between 2013 and 2016 by red circles. Running 31-year average values are indicated by the yellow squares. The smoothed curves represent 3rd order polynomial fits to the data over three 97-year windows: 1901–1997 (*black*), 1916–2012 (*blue*) and 1920–2016 (*red*): **a** Globe and **b** Pacific. In both cases, the six panels show the effect of applying different polynomial fits to the data

warming up to the present, despite the relatively stagnant period.

The raw data, including the 31-year average values, are well represented by 3rd order polynomial fits. Three such fits to the raw data are also shown based on overlapping 97-year windows (1901–1997, 1916–2012, and 1920–2016) and unlike running averages, provide estimates of the magnitude of background trends at the beginning and end of each window. The effect of applying different polynomial fits to the data is indicated by the panels in the figure, which show the estimated background trends at the end of each window. The higher order fits reflect the relatively shorter term fluctuations in the data compared to the lower order fits. A 3rd order polynomial provides a close fit to the data since (a) 85 % of the variance in the 1921–2016 data is explained, (b) this percentage is more than explained by a 2nd order fit (80 %) yet not much less than explained by the higher order fits. Importantly, the 3rd order polynomial fits yield almost identical background trends, indicating that they capture a robust long-term feature of the data. In contrast, the trends associated with the higher order fits tend to differ, being more sensitive to shorter-term fluctuations including hiatus periods. As a consequence, the data can be described as comprising a robust, ever-increasing long-term warming trend that is independent of shorter-term fluctuations at the end of the series.

A different picture emerges for Pacific average temperatures (Fig. 1b), as the 31-year average temperature anomalies (yellow squares) and a 3rd order polynomial reflect the existence of a slow-down over the period 1990–2000. All fits to the data over the windows reveal that, apart from the linear fits, the background trends all differ considerably in response to the values after 1997. In each case, the fits based on the early 1901–1997 window indicate relatively moderate to strong warming trend, the 1916–1997 window yields far less warming and even cooling, while the recent 1921–2016 window indicates a return to moderate warming. These features reflect a distinct, but temporary, hiatus in the long-term warming trend in the Pacific.

We subtract the raw global values from the 3rd order polynomial fit to the full data set (1900–2016) to obtain the residual or detrended values (black, Fig. 2). These represent the effects of internal variability at relatively shorter

time scales including contributions from the IPO, El Niño/La Niña events and any other external sources including volcanic eruptions. The IPO can only explain about 16 % of the residual variance in global temperatures, mainly because of regional cancellations (Chen et al. 2008). The IPO-based values (using linear regression) are indicated by the green and indicates that it can account for no more than about -0.03 °C to the slowdown in observed temperatures after 1998. The NINO34 index indicates that this contribution comprises five La Niña events (1999, 2000, 2007, 2008 and 2011) and only a single El Niño event (2002) between 1999 and 2014.

4 Comparing model simulations with observations

Figure 3 compares both the observed global and Pacific average surface temperature anomalies (relative to 1961–1990) over the period 1900–2016 with MMM values from the CMIP5 simulations using the RCP4.5 emissions scenario. Note that for the time period considered, there is very little difference between the MMM values from either the RCP8.5, RCP4.5 or RCP2.6 simulations. Third order polynomial fits to the data over the period are shown in order to better compare the long-term trends in the data sets. In both Figures the model values and the observations behave similarly after about 1910, up until about 1990, but appear to diverge thereafter. By 2016 the model ensemble values are about $+0.2$ °C and $+0.3$ °C warmer than the observations for the globe and the Pacific respectively. In particular, the MMM global temperatures consistently exceed the observed values every year after 1998 and, for Pacific temperatures, every year after 1993. This period is also unusual for the fact that the observed global temperature was lower than the temperature in all ensemble members, for all years between 2011 and 2014, except in only 10 out of a possible 332 instances.

The probability that the MMM average value for 1986–2016 comes from the same population as the observations is only 0.14 according to a simple *t* test. However, these differences are dependent on the baseline period we use to define the anomalies (Hawkins and Sutton 2015). For example, these differences would disappear if we used the 1986–2016 period to define the anomalies in both the models and observations, but at the expense of making the models appear unrealistically cool for much of the twentieth century. If we take the period 1900–1985 as the baseline period, then the model values based on the 1961–1990 period are warmer by $+0.03$ °C. By subtracting this amount from each model value we can make the 1900–1985 average value match that of the observations. While this reduces the differences at the end of the time series, it does not greatly affect the divergence between them. For

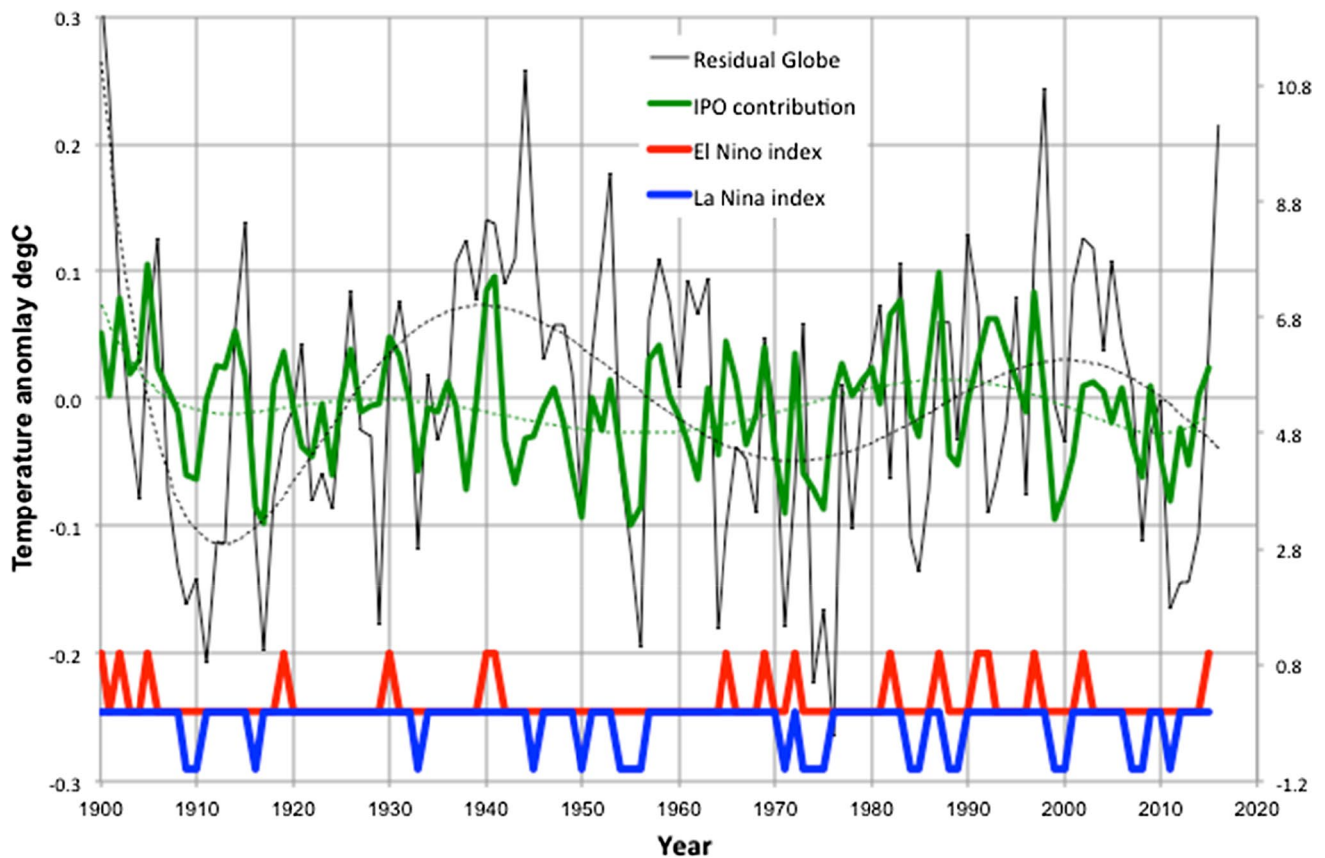


Fig. 2 Detrended Pacific temperature anomalies (*black*) versus anomalies associated with the IPO (*green*). El Niño (*red*) and La Niña (*blue*) years based on NINO34 temperatures are indicated

the 1986–2016 period, the resultant probability increases to 0.34, but not enough to indicate that the MMM values as a whole are consistent with the observations, despite the record warm years in the 2015 and 2016. The choice of baseline does not affect the interpretation of the trends in the data.

The issue here is that we expect that, at a certain time scale, the effects of greenhouse-related climate change to dominate, and any other sources of variability will be represented at shorter time scales. The fact that we find the observed data is well represented by a 3rd order polynomial (a form of low pass filter), and that the implied trends at the end of successive 100-year windows are almost identical, suggests that we can identify just such a signal and that other sources of variability occur on time scales much less than about 50 years. i.e. hiatus periods of up to 20 years represent variability not associated with a robust background warming trend. Therefore, using the same “filter” on model-based data allows a comparison to be made of the actual and simulated warming signal. Furthermore, the MMM global warming trend is effectively unaffected by any simulated internal variability since this tends to cancel out across the different simulations. As a consequence,

the difference between the observed temperatures and the MMM values of about $+0.2\text{ }^{\circ}\text{C}$ is unlikely to represent the effects of internal variability. As already noted, the IPO can only explain about $0.03\text{ }^{\circ}\text{C}$ of the detrended global values (Fig. 2).

A short-term excess of La Niña events over El Niño events between 2003 and 2014 contributed to the observed Pacific hiatus. However, over the longer period 1986–2016 the excess was only 1 (Fig. 2). Over this same period, based on simulated Pacific temperatures, the average excess of La Niña-like events over El Niño-like events across all 83 simulations was -0.66 (i.e. slightly more El Niño-like events than La Niña-like events) but not significant given that the standard deviation of the excess values across all simulations is 2.9. These numbers indicate that changes in the frequency of El Niño/La Niña events can affect short-term trends but contribute relatively little to any changes in long-term trends. Even if we assume that an excess of La Niña events can explain a similar amount as the IPO (which strictly cannot be done because these variables are closely related, Power et al. 2005; Newman et al. 2003) the total contribution would still only be about $0.06\text{ }^{\circ}\text{C}$.

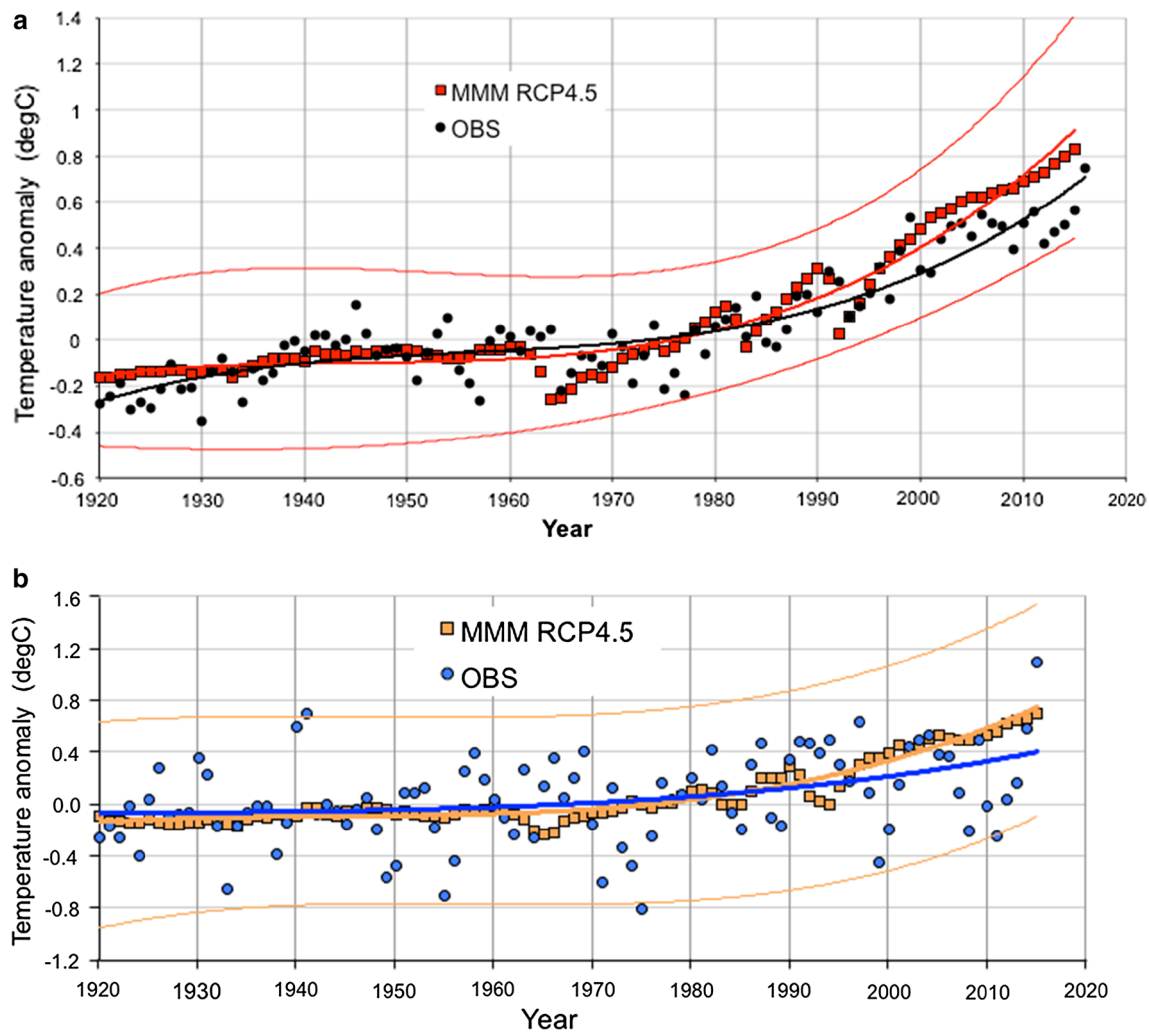


Fig. 3 Annual average surface temperatures 1921–2016 from observations and multi-model mean values from CMIP5 (RCP4.5) simulations. The temperatures are all expressed as anomalies relative to 1961–1990. The smoothed curves are based on 3rd order polynomial fits to the data over the period 1920–2016. The upper and lower

smooth curves correspond to the multi-model minimum and maximum values respectively. **a** Observed global values (*black*), model values (*red squares*). **b** Pacific values (*blue circles*), model values (*orange squares*). Note differences in scale

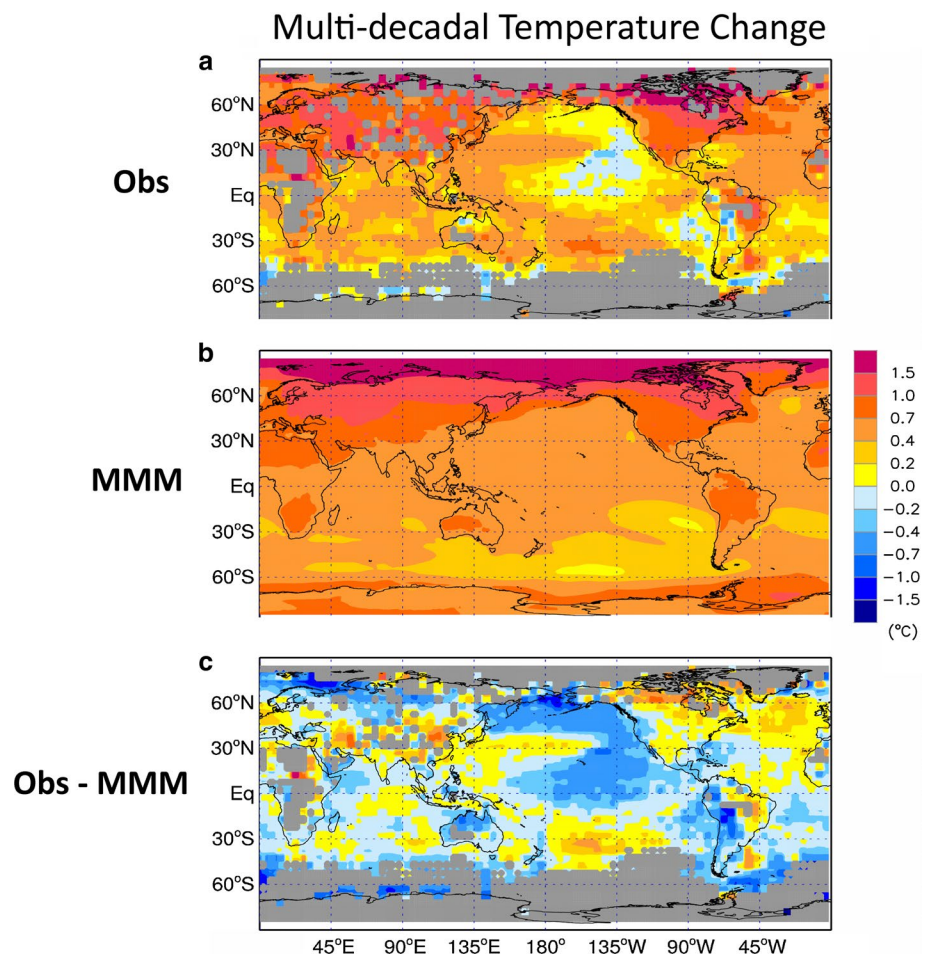
The Atlantic Multi-decadal Oscillation (AMO) represents North Atlantic sea surface temperature variability and appears to represent a source of variability in global average temperatures independent of the Pacific (Chylek et al. 2014). However, observations suggest that it would have contributed relatively warm values over recent years (see Fig. 4c). The observed AMO cannot, therefore, help explain why the observed global temperatures are cooler than the MMM values.

In the case of the Pacific, the situation is different since the background warming trend does appear affected by a hiatus (Fig. 1b). At the end of the time series the model values again appear too warm by about $+0.3^{\circ}\text{C}$. If we use the 1900–1985 baseline, the difference is slightly less ($+0.28^{\circ}\text{C}$). The IPO explains about 68 % of

the variance in the Pacific time-series but can explain no more than about 0.15°C during the hiatus period. This is about half the difference between the observed and MMM values at the end of the time series indicating that, even for this region, the models, as a whole, appear too warm. It is worth noting that this difference is persistent, and often exacerbated, by the choice of different baseline periods.

Finally, while the IPO makes a contribution to the recent multi-decadal change in Pacific temperature (Sect. 3.1; Fig. 10), the multi-decadal change in the IPO does not appear large in terms of its own historical record (“see Appendix”). This does not support the hypothesis (see, e.g., Kociuba and Power 2015) that the contrast between observed and modeled multi-decadal variability can be

Fig. 4 Spatial pattern of multi-decadal temperature change (from 1961–1990 to 1998–2012, HADCRUT4) from **a** observations (HADCRUT4), **b** models (MMM) and **c** the difference between them. (Observations-MMM)



accounted for by an extreme natural excursion of the IPO in the real world.

5 Pacific variability

Figure 4a shows the global pattern of multi-decadal temperature change (i.e., temperature change from 1961–1990 to 1998–2012). Although the changes depicted are dominated by the long-term warming trend, it also reflects a contribution by the IPO consistent with earlier research (e.g., Trenberth and Fasullo 2013; England et al. 2014; Meehl and Arblaster 2012; Goddard 2014; Kociuba and Power 2015). Figure 4b shows the corresponding MMM temperature anomaly pattern. It appears relatively featureless over the entire Pacific Ocean when compared with the observed pattern (Fig. 4a). This is expected since averaging across model simulations largely filters out any internally generated features, including the impact of the IPO. The difference between the two patterns (Fig. 4c) therefore emphasizes the contribution of the real world IPO, and other sources of internal variability, to the model-to-observed contrast in the multi-decadal temperature change.

In fact, if models and the forcing applied to them were perfect, and observational error is ignored, then non-zero values in Fig. 4c would – by construction – arise entirely from interdecadal climate variability.

Figure 4c indicates that the (cool) difference between the observed and MMM values exceeds 0.7 °C in the central Pacific. Is internal variability large enough so we can reasonably expect it to account for such a value? In order to address this question, let us assume – for the time being – that modelled internal multi-decadal variability has a realistic magnitude. Figure 5a shows the spatial structure of the standard deviation (SD) of multi-decadal variability arising from internal climate variability alone, as it appears in pre-industrial simulations. If we suppose that the models and the forcing applied to them are perfect, then the population standard deviation is given by $SD^* = SD(1 + 1/N)^{1/2}$, where N is the number of simulations. Defining $D2012$ as the difference between the observed temperature change over the period 1998–2012 (relative to 1961–1990) and the MMM value over the same period, then the ratio $D2012/SD^*$ provides a measure of the role of internal climate variability. For example, if $D2012/SD^* = 1.0$ in a particular location, then the difference is equivalent to 1

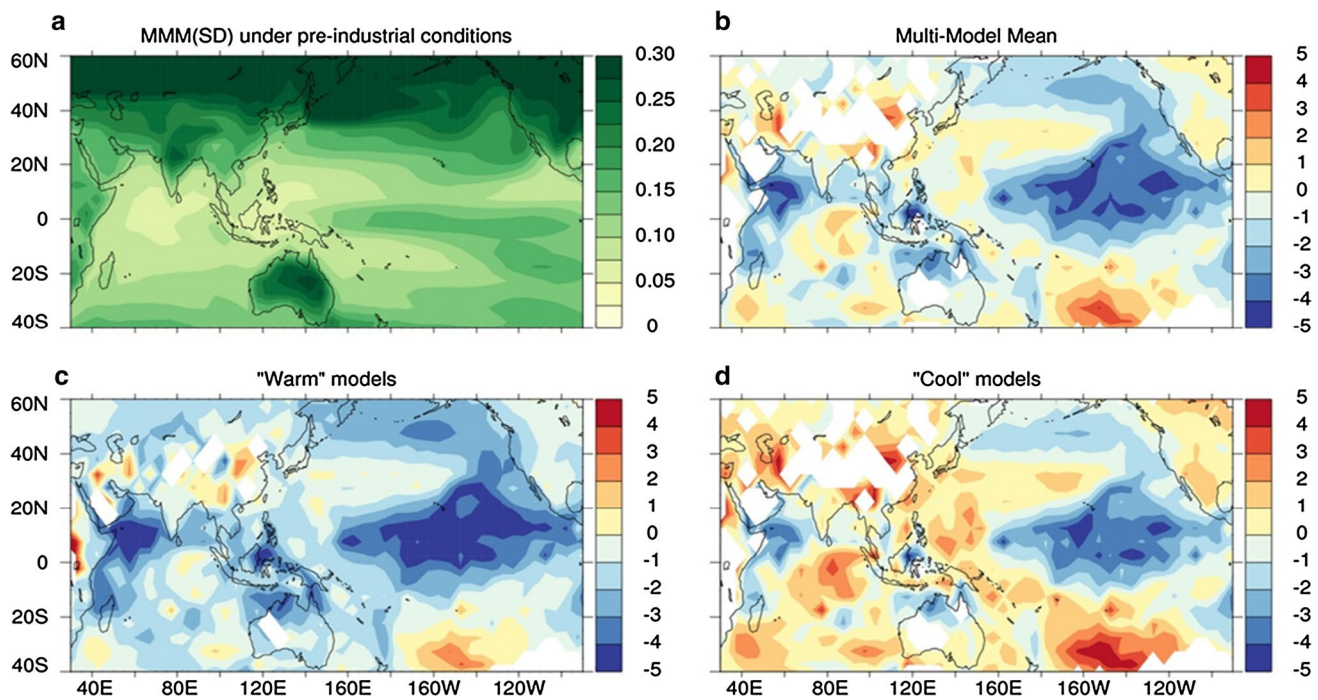


Fig. 5 Estimates of the contribution of internal climate variability to recent multi-decadal temperature change over the Indo-Pacific **a** multi-model mean (MMM) of the standard deviation (SD) of $\Delta T = T_{15}(t) - T_{30}(t)$ (°C) that can arise from internal climate variability alone, and the (spatially varying) ratio $D(2012)/SD^*$, where SD^* is the standard deviation of variability that can arise in $D(2012)$ from internal climate variability alone, under the assumption that the models and the forcing applied to them are perfect. **b** All models, **c** ten warmest models only and **d** ten coldest models only, are used to calculate both D and SD^* . Here $D(2012) = \Delta T_{obs}(2012) - \text{MMM}[\Delta T_{models}(2012)]$, T_{15} is a 15 years average ending in year t , and T_{30}

is an average over a 30 years period ending 22 years prior to t . This statistic is chosen to match the averaging lengths and gaps between the periods used in Fig. 6 (i.e. the 15-year period 1998–2012 and the preceding thirty-year period 1961–1990). The estimates presented are based on the internal climate variability evident in pre-industrial simulations (see Supplementary Table 1 for a list of the simulations analysed). Under the assumption that the models and the forcing applied to them are perfect, the standard deviation of the difference, D , between ΔT_{obs} and the MMM of ΔT_{models} is given by $SD^* = SD(1 + 1/N^2)^{1/2}$, where N is the number of models used in the calculation of the MMM

SD^* , indicating that it is easily accounted for by internal variability.

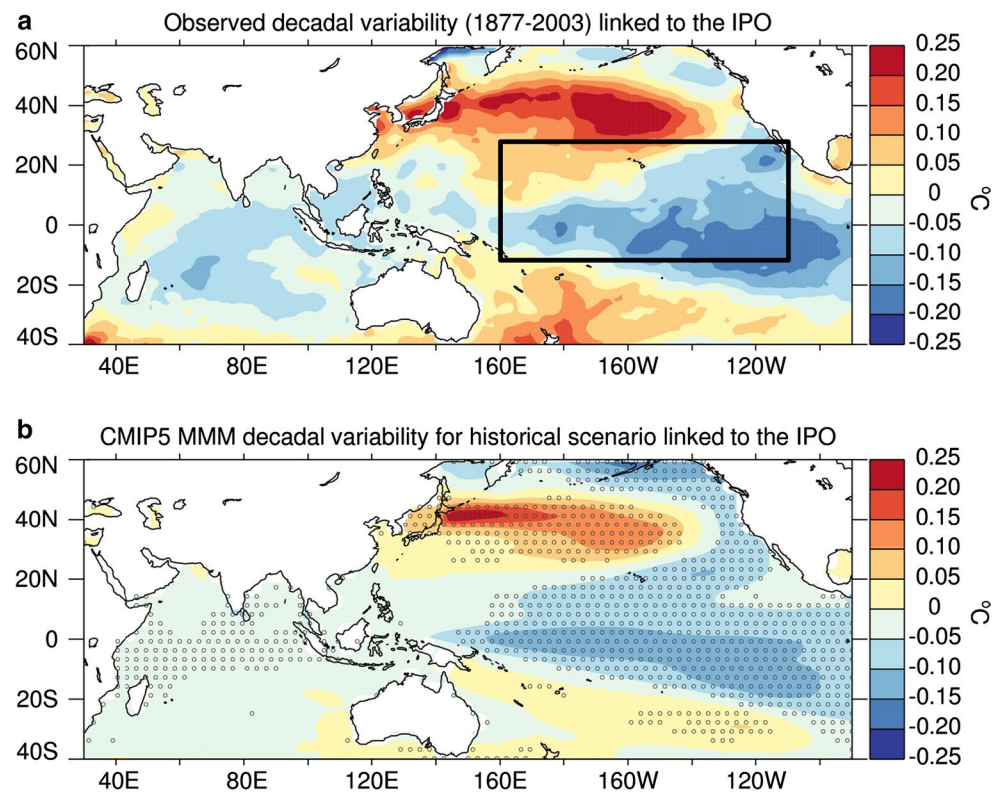
Figure 5b shows the pattern of $D2012/SD^*$ for all simulations, while Fig. 5c, d show the patterns for the ten “warmest” (based on $D2012$) and the ten “coolest” simulations respectively. In all three cases (in Fig. 5b–d) the magnitude of $D2012/SD^*$ over the Pacific Ocean actually exceeds 4, indicating that the differences are extremely large compared with modelled internal multi-decadal variability. In fact the warmest models exhibit ratios exceeding five in some locations. The model-to-observed contrast in $D2012$ is therefore very unlikely—especially in the case of the warmest models—to be caused by internal climate variability alone.

This conclusion is partially based on the assumption that models simulate a realistic level of internal climate multi-decadal variability. But do they? Given the importance of the IPO to multi-decadal variability in the Pacific, let us begin answering this question by examining how well the models simulate the IPO. IPO patterns were diagnosed in

each of the model simulations. The pattern matching correlation coefficient with the observed IPO ranges from 0.3 to about 0.75 (c.f. Brown et al. 2014). The MMM IPO pattern (Fig. 6b) resembles the observed pattern (Fig. 6a) but the modelled temperature excursions depicted are not as large. This is consistent with the findings of Kociuba and Power (2015). They concluded that the difficulty models have in simulating the recent multi-decadal strengthening of the Walker circulation may be partially due to a systematic underestimate of internal interdecadal variability in the strength of the Walker circulation. It also consistent with the results of England et al. (2014), who drew similar conclusions in relation to Pacific trade wind strength. This is also consistent with the fact that models tend to underestimate autocorrelation arising from El Niño–Southern Oscillation, which can make it difficult for models to generate enough decadal and longer-term IPO-related variability (Kociuba and Power 2015).

While the analysis above suggests that modelled IPO variability might be too weak, the observational record is

Fig. 6 Standard deviation of SST variability (1900–2005) linked to the Interdecadal Pacific Oscillation (IPO): **a** observations (Rayner et al. 2003), **b** multi-model mean (MMM). Stippling indicates that over 70 % of models have the same sign as the multi-model mean, which exceed the 99 % statistical significance level under the assumption of model independence (Power et al. 2012)



relatively brief for this purpose, and so our level of confidence in the conclusion that models underestimate the level of multi-decadal variability in the Pacific is not high. Furthermore, the modelled internal multi-decadal climate variability would have to be far weaker than real-world internal interdecadal climate variability in order for this to account for the model-to-observed contrasts over the Pacific (in Figs. 4c, 5b–d). For example, if the standard deviation is doubled to account for this possible model deficiency, then the ratio (Fig. 5b–d) still exceeds 2.5. A naturally occurring excursion of the real-world IPO would have to be very large in historical terms. But this does not appear to be the case, as the recent multi-decadal excursion of the IPO index (“see Appendix”) is not unusual in terms of the IPO’s own historical record.

We therefore conclude that it is highly unlikely that Pacific Ocean internal variability (even if allowance is made for possibly weaker than observed simulated variability) alone can account for the resulting contrast between the observations and the model simulations (in Figs. 4c, 5b–d; especially Fig. 5c—the “warmest” models).

Finally, note that while Huber and Knutti (2014) argue that ENSO variability could account for a 15-year cooling trend of -0.06 °C, we find little evidence that IPO and El Niño/La Niña events can explain the apparently robust MMM warm bias in multi-decadal temperature changes.

6 Model temperature projections

Figure 7 compares observed and individual model values (15-year trailing averages of anomalies relative to 1961–1990) for both the globe and the Pacific. The observed global value in 2012 (Fig. 7a) is lower than 90 % of the individual values, while the corresponding Pacific value (Fig. 7b) is lower than *all* the model simulations analysed. Also shown are the evolution of the ten warmest (based on D(2012)) simulations (red lines) and the ten coolest simulations (blue lines). The plot indicates that the projected global values at the end of the twenty-first century partly depend on the values at the start of the twenty-first century, since the red and blue lines tend to separate. On the other hand, there is no evidence for a comparable separation over the Pacific.

Figure 7 also indicates that model-to-model variability in multi-decadal temperature change in the Pacific (between 1961–1990 and 1998–2012) is largely driven by internal variability, whereas model-to-model variability in recent multi-decadal *global* temperature change is partially driven by model-to-model differences in the response of models to external forcing.

Figure 8 provides model results from three different twenty-first century scenarios in which GHG emissions and concentrations are highest under RCP8.5, lower under RCP4.5 and lowest under RCP2.6. It can be seen that the

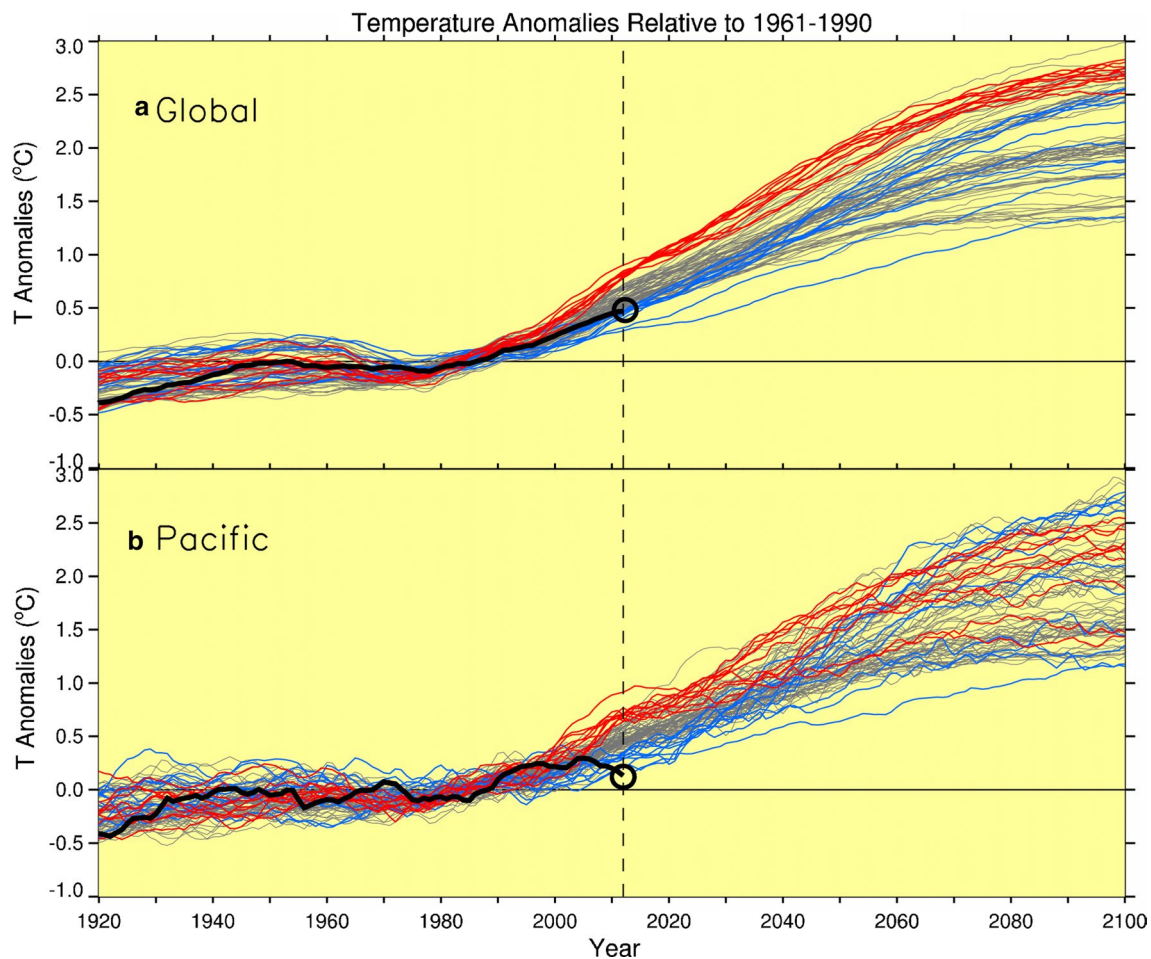


Fig. 7 Trailing 15-year average surface air temperature anomalies, relative to preceding 30-year averages. For example, the values in 2012 represents the multi-decadal difference 1998–2012 relative to 1961–1990. Observed (HadCRUT4, black) and model values (1920–

2100, red, blue and grey): **a** Global and **b** Pacific. The red lines correspond to the ten warmest values at 2012, the blue lines to the ten coolest, and the grey lines to the remainder

simulations that yield relatively warm changes over the past half-century (red lines) tend to produce larger projected changes. Similarly, simulations that yield relatively cool temperature changes over the past half-century (blue lines) tend to exhibit smaller projected changes. This finding is independent of the baseline period chosen. The contrast between late twenty-first century projections from the models exhibited the greatest change over the past half-century and the models exhibiting the least change over the past half-century is clearest under both RCP8.5 (Fig. 8a) and RCP4.5 (Fig. 8b).

The link between recent multi-decadal temperature change and projected temperature change is further illustrated in Fig. 9, which shows the projected global average temperature change for the period 2086–2100 versus the simulated multi-decadal temperature change for the period 1998–2012 (both changes relative to 1961–1990). In each case, the relatively warm present-day models yield

relatively warm projections and the strength of this effect increases with the higher emissions scenario. For RCP 8.5 (Fig. 9a) the correlation coefficient (r) associated with the line-of-best-fit is $+0.45$, and is statistically significant with a p value of $.005$.

Note that this does not necessarily mean that the warming rate between 1998–2012 and 2086–2100 is necessarily greater for the warmer models. Suppose we define “cool models” as models that simulate recent warming less than 0.35 °C, and “warm models” as models that simulate recent warming greater than 0.55 °C. Then under RCP8.5 the “warm models” warm more than the “cool models” (4.15 °C compared with 3.71 °C), whereas the warming is similar for both “warm models” and “cool models” under RCP4.5 (1.98 and 1.96 °C, respectively), and under RCP2.6 the “warm models” actually warm less than the “cool models” (1.03 and 1.23 °C, respectively) between the two periods (i.e. 1998–2012 and 2086–2100).

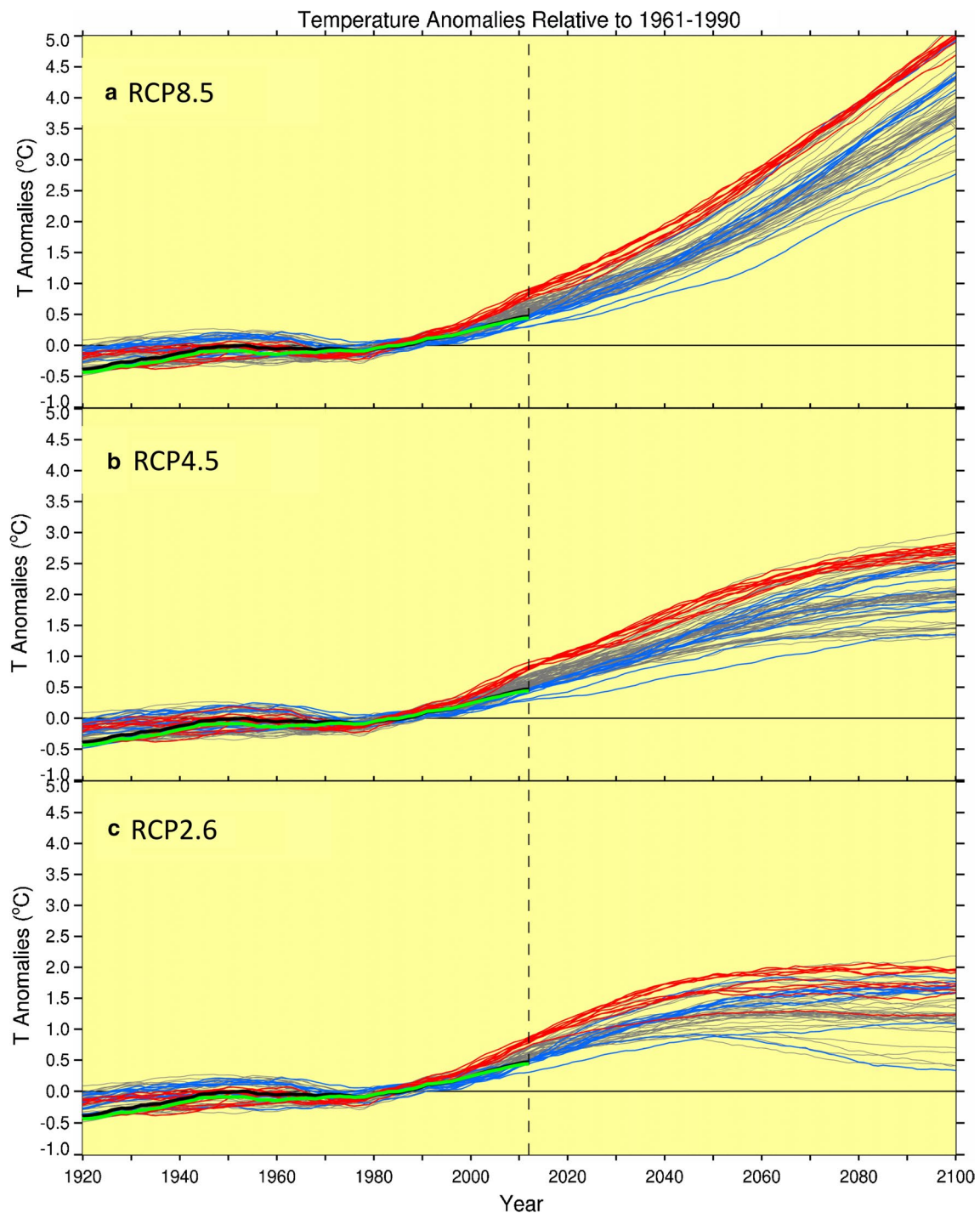


Fig. 8 Trailing 15-year global average surface air temperatures. Observed (HadCRUT4, *black*) and model values (1920–2100, *red*, *blue* and *grey*): **a** RCP8.5, **b** RCP4.5 and **c** RCP2.6. The *red lines* cor-

respond to the ten warmest MMM values at 2012, the *blue lines* to the ten coolest, and the *grey lines* to the remainder

We might therefore conclude that model-to-model deviations in the magnitude of global warming over the past-half century does not seem to provide simple or clear guidance for twenty-first century temperature change for the same models. These results appear broadly consistent

with those of England et al. (2015), who showed that model-to-model deviations on recent, much shorter (15-year) trends do not provide any guidance on model-to-model deviations in the magnitude of late twenty-first century warming.

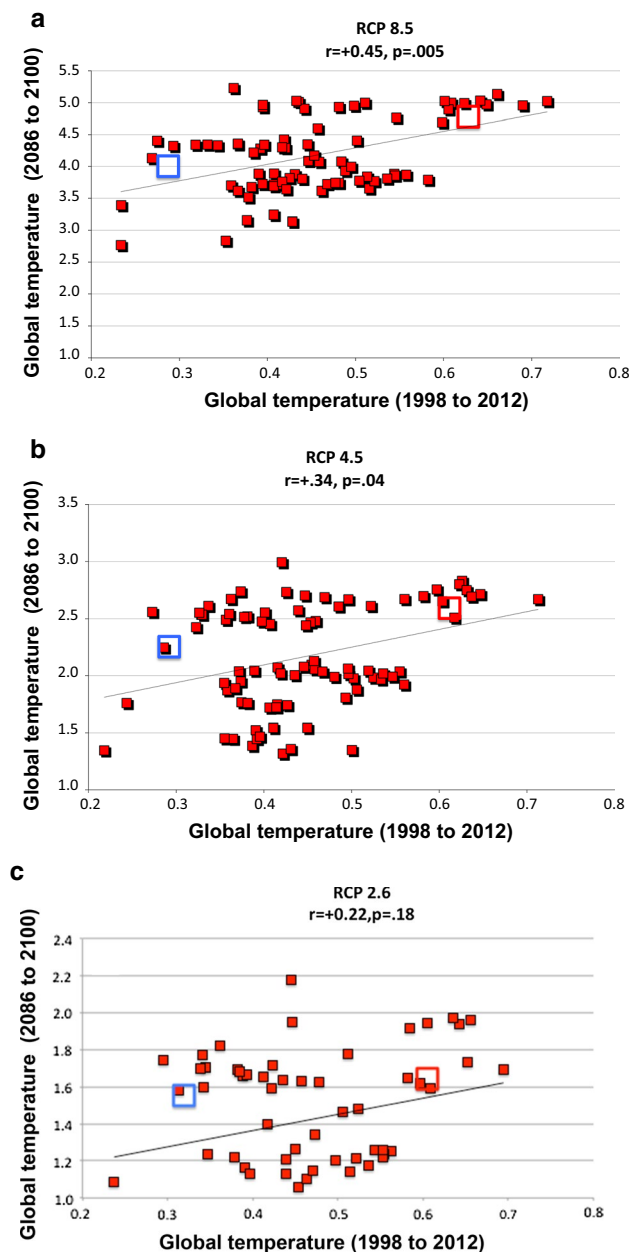


Fig. 9 Projected global average temperature change for the end of the century 15-year period (2071–2100) (relative to 1961–1990) versus model average multi-decadal temperature change from 1961–1990 to 1998–2012, for three different emissions scenarios. The *blue and red squares* indicate the average values for “cool” and “warm” models, respectively. Lines-of-best-fit, and associated correlation coefficients and *p* values, are shown. **a** RCP8.5, **b** RCP4.5 and **c** RCP2.6

However, under RCP8.5 the “warm models” do tend to warm more than the “cool models” between 1998–2012 and 2086–2100, and there is a (very marginally) larger warming of the “warm models” under RCP4.5. This leaves open the possibility that some of these models might overestimate the warming response to the imposition of

greenhouse gas increases. But clearly, model-to-model contrasts in the sensitivity to GHG increases alone cannot fully explain the results.

It may be that model-to-model contrasts in the response to *both* (a) GHG increases and (b) sulphate aerosol changes may need to be taken into account in order to understand the results. This possibility can arise if “cool models” tend to cool more than the other models in response to sulphate aerosol increases and if “warm models” tend to warm more than the other models in response to GHG increases. Factors (a) and (b) could then explain the results if: (a) dominates the contrasts in the twenty-first century warming rates between “warm” and “cool models” under RCP8.5 (which has the largest GHG increases); (b) dominates under RCP2.6 (which has the smallest GHG increases); neither dominates under RCP4.5 (which has GHG increases that lie between those of RCP2.6 and RCP8.5).

Finally note that since the modelled multi-decadal changes over the past half-century as a whole appear too high, most likely because of issues associated with forcings and/or the sensitivity to GHGs and other anthropogenic forcing (Flato et al. 2013; Kirtman et al. 2013), it follows that the projections might also be too high.

7 Summary and discussion

The vast majority of the numerous earlier papers related to the so-called global warming hiatus (e.g. Fyfe et al. 2013; Fyfe and Gillett 2014; Flato et al. 2013; Hawkins et al. 2014; Kosaka and Xie 2013; see Introduction for additional references) focussed on relatively short-period (e.g. 15 year) trends. Here we primarily focus on much longer-term (multi-decadal changes). A key example of a multi-decadal change is a change from a thirty-year reference period (1961–1990) to a later 15-year period (e.g. 1998–2012). Such metrics measure changes over approximately half a century. This very different focus means that earlier conclusions drawn on the basis of analyses examining much shorter-term change (e.g. trends over 15 year periods) do not necessarily apply.

We examined decadal, multi-decadal and longer-term changes in global and Pacific temperatures using data up to near-present, and compared these to simulated changes over the same period. We identified large and important model-to-observed contrasts, and we examined the implications the contrasts have for projections of global and Pacific temperature over the remainder of the twenty-first century. In order to do this we also examined the ability of models to simulate multi-decadal climate variability and the degree to which internal, multi-decadal climate variability can account for the model-to-observed contrasts.

Table 1 Details of MMM temperatures anomalies ($^{\circ}\text{C}$) with respect to 1961–1990. The observed (HadCRUT4) value for 1998–2012 is $+0.39^{\circ}\text{C}$

Ensemble (N)	MMM global average temperature 1998–2012	MMM global average temperature 2086–2100
RCP2.6 (60)	+0.46	+1.4
RCP4.5 (83)	+0.45	+2.2
RCP8.5 (76)	+0.45	+4.2

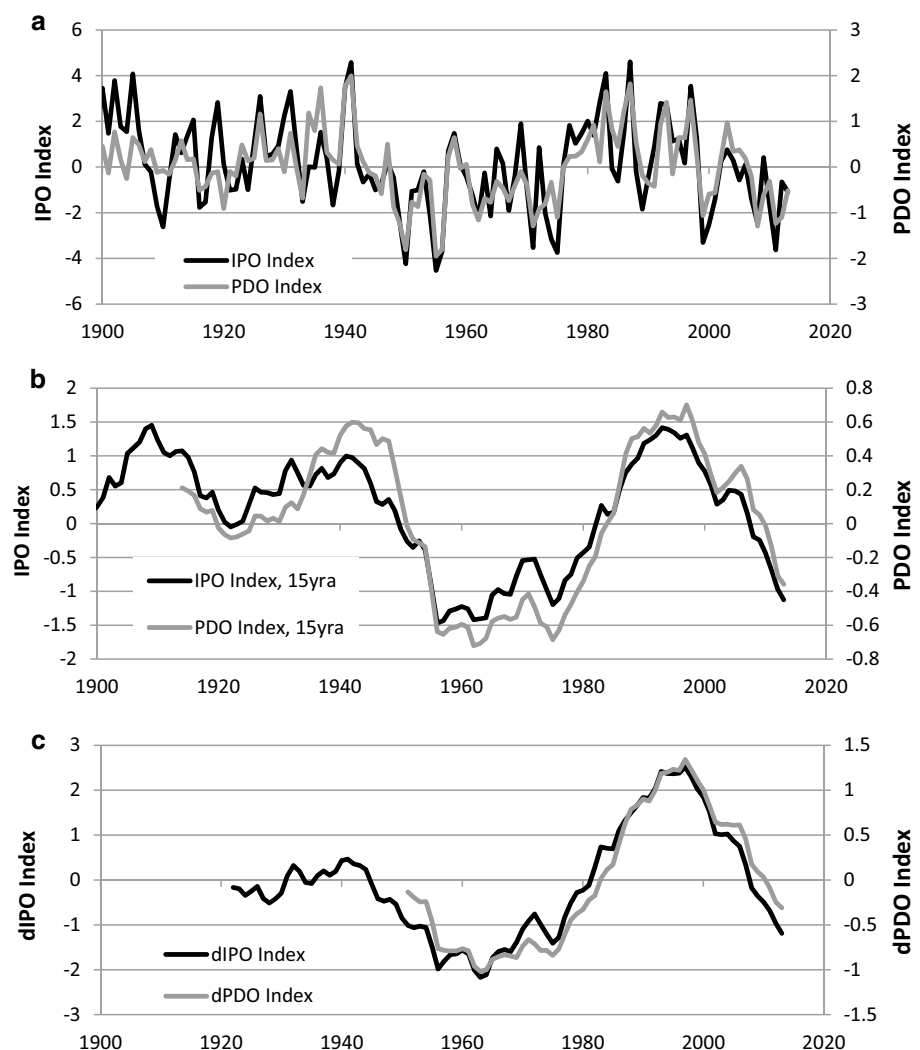
A key finding is that, in *all* of the CMIP5 model simulations analysed, for the period 1998–2012 relative to 1961–1990, the Pacific warms more on multi-decadal time-scales than do the observations. Furthermore, at the global scale, 90 % of the simulations exhibit greater than observed multi-decadal surface warming.

The more moderate observed warming is due in part to cooling by the IPO and from an increase in the frequency of La Niña events. However, we find that the magnitude of the contrast between the observed and simulated multi-decadal

changes in Pacific temperature are so large that it is highly unlikely that Pacific Ocean internal variability alone, even if allowance is made for possibly weaker than observed simulated variability, can account for the contrast. This is especially true for the contrast between observations and the models exhibiting the greatest multi-decadal warming. We also find that, while the IPO makes a contribution to the recent multi-decadal change in Pacific temperature, the corresponding multi-decadal change in IPO indices is not large in terms of its own historical record. This does not support the hypothesis (see, e.g., Kociuba and Power 2015) that the contrast between observed and modeled multi-decadal variability over the past half-century can be accounted for by an extreme, natural excursion of the IPO in the real world.

Together, this indicates that imperfections in the models or the forcing applied to them are partially responsible for the model-to-observed contrast in multi-decadal temperature change over the past half-century. To be relevant the imperfections must result in too little cooling during 1998–2012 relative to the reference period (1961–1990), or

Fig. 10 Time series associated with the observed IPO and PDO. **a** Annual indices for the IPO and PDO, **b** 15 years running averages of the indices IPOI_{15} and PDOI_{15} , **c** and the differences $\text{IPOI}_{15}(t) - \text{IPOI}_{30}(t)$ and $\text{PDOI}_{15}(t) - \text{PDOI}_{30}(t)$. Here IPOI_{15} is a 15 years average of the annual IPO index ending in year t , and IPOI_{30} is an average of the annual IPO index over the 30y period ending 22 years prior to t . This particular statistic is chosen to match the averaging lengths and gaps between the periods used in Fig. 1 (i.e. the 15-year period 1998–2012 and the preceding thirty-year period 1961–1990). These statistics help to provide an indication of what the IPO and PDO contributed to e.g. $\Delta T_{\text{obs}}(2012) = T_{\text{obs}}(1998-2012) - T_{\text{obs}}(1961-1990)$



too little warming during the reference period. This could include imperfections in the forcing used in the simulations over the past half-century. Candidates for errors in forcing include: an underestimate of stratospheric aerosol concentrations; reduced solar output; lower levels of water vapour in the upper atmosphere; sulphate aerosols; and the impact of GHGs on warming in the tropical Pacific (Andersson et al. 2015; Clement et al. 1996; Flato et al. 2013; Huber and Knutti 2014; Santer et al. 2014; Schmidt et al. 2014) in the more recent period (e.g. 1998–2016). Huber and Knutti (2014) estimate that underestimating stratospheric and solar factors could have contributed to a cooling trend of $-0.07\text{ }^{\circ}\text{C}$ over 15 years but appears insufficient to account to about for the approximately $0.2\text{ }^{\circ}\text{C}$ difference at the end of the time series. Another candidate for examining the model-to-observed contrast is that some of the models—especially those models that exhibit the greatest multi-decadal warming to date—may overestimate the warming response to the imposition of greenhouse gas increases.

We showed, however, that model-to-model contrasts in the sensitivity to GHG increases alone do not fully explain the model-to-model contrasts in projected global temperature increases from 1998–2102 to 2086–2100. Instead we showed that competing, model-to-model contrasts in the response to both (a) GHG increases and (b) sulphate aerosol might explain the results if we assume that: (a) dominates the contrasts in the twenty-first century warming rates between “warm models” (i.e., the models that warmed the most over the past half-century) and “cool models” (i.e., the models that warmed the least over the past half-century) under RCP8.5 (which has the largest GHG increases among the RCP scenarios); (b) dominates under RCP2.6 (which has the smallest GHG increases); neither (a) nor (b) dominate under RCP4.5 (which has GHG increases that lie between those of RCP2.6 and RCP8.5). This could then explain: why “warm models” (i.e. the models that warmed the most over the past half-century) warm more than “cool models” under RCP8.5 from 1998–2012 to 2086–2100; why the temperature changes between 1998–2012 and 2086–2100 are similar in the “warm models” and “cool models” under RCP4.5; and why the “cool models” warm more than the “warm models” do, between 1998–2012 and 2086–2100, under RCP2.6.

Further substantial warming over the twenty-first century is nevertheless projected in all the models under business as usual emission scenarios (RCP4.5 and RCP8.5; Figs. 1, 6 respectively). If attention is restricted to models with more accurate simulations of recent multi-decadal temperature change then the resulting model ensemble has far fewer members exhibiting the highest warming levels in the late twenty-first century. The warming nevertheless remains much larger than warming to date, under both the RCP4.5 and RCP8.5 scenarios (Table 1).

Finally, we found that observations up to 2016 show no evidence of a hiatus in the background global warming trend. This is consistent with the findings from other studies (Karl et al. 2015, Lewandowsky et al. 2015, 2016) that stress the importance of multi-decadal time scales rather than short-term trends (e.g. Fyfe et al. 2013; Flato et al. 2013; Hawkins et al. 2014; Kosaka and Xie 2013). However, the data do indicate a temporary hiatus in the warming trend for the Pacific, due in part to cooling by the IPO associated with a short-term excess of La Nina events between 2003 and 2014.

Acknowledgments This work is supported by the Pacific-Australia Climate Change Science and Adaptation Planning Program (PAC-CSAP), the Australian Climate Change Science Program (ACCSPP), and the National Environmental Science Program (NESP). We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Lawson Hanson, Harvey Ye and Aurel Moise made the CMIP5 data available to us. Rob Colman and Brad Murphy provided helpful comments on an earlier version of the paper. Thanks also to Dr. G. Meehl and an anonymous referee for excellent and constructive suggestions for improving the paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Impact of the IPO on recent, observed multi-decadal climate change

In this appendix we use alternative measures of the IPO to see, first, if the IPO made a contribution to recent observed temperature change, and second, if the recent observed change in the IPO index is unusually large in terms of its own historical variability. The index for the observed IPO from the UKMO and the index for the closely related Pacific Decadal Oscillation are given in Fig. 10a. The two time-series are closely related (Fig. 10a; $r = 0.78$), especially on interdecadal time-scales (Fig. 10b; $r = 0.94$). This is consistent with the view that the PDO is the North Pacific expression of the near-global-scale IPO. Both indices are, as expected (England et al. 2014; Meehl and Arblaster 2012; Goddard 2014), generally in a negative phase during 1998–2012 (Fig. 10b). However, the impact on recent multi-decadal temperature change (i.e. from 1961–1990

to 1998–2012 for example) will depend on what the IPO did during both 1998–2012 and the reference period 1961–1990. The difference in the IPO and PDO indices between these two periods (Fig. 10c) is negative, indicating that they both acted to reduce multi-decadal warming, again as expected. However, the magnitudes of the multi-decadal changes in the indices are very modest compared with some earlier periods (Fig. 10c). This does not support the hypothesis that recent multi-decadal change in the IPO is extreme.

References

- Andersson SM, Martinsson BG, Vernier J-P, Friberg J, Brenninkmeijer CAM, Hermann M, van Velthoven PFJ, Zahn A (2015) Significant radiative impact of volcanic aerosol in the lowermost stratosphere. *Nat Commun* 6:7692. doi:[10.1038/ncomms8692](https://doi.org/10.1038/ncomms8692)
- Bindoff NL, Stott PA, AchutaRao KM, Allen MR, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov II, Overland J, Perlwitz J, Sebbari R, Zhang X (2013) Detection and attribution of climate change: from global to regional. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Brown PT, Li W, Li L, Ming Y (2014) Top-of-atmosphere radiative contribution to unforced decadal global temperature variability in climate models. *Geophys Res Lett*. doi:[10.1002/2014GL060625](https://doi.org/10.1002/2014GL060625)
- Brown PT, Li W, Xie SP (2015) Regions of significant influence on unforced global mean surface temperature variability in climate models. *J Geophys Res Atmos*. doi:[10.1002/2014JD022576](https://doi.org/10.1002/2014JD022576)
- Chen J, Del Genio AD, Carlson BE, Bosilovich MG (2008) The spatiotemporal structure of twentieth-century climate variations in observations and reanalyses. Part II: Pacific pan-decadal variability. *J Clim* 21:2634–2650
- Chylek P, Klett JD, Lesins G, Dubey MK, Hengartner N (2014) The Atlantic Multidecadal Oscillation as a dominant factor of oceanic influence on climate. *Geophys Res Lett* 41:1689–1697. doi:[10.1002/2014GL059274](https://doi.org/10.1002/2014GL059274)
- Clement AC, Seager R, Cane MA, Zebiak SE (1996) An ocean dynamical thermostat. *J Clim* 9:2190–2196
- Cowtan K, Way RG (2014) Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Q J Roy Meteorol Soc* 140(683):1935–1944. doi:[10.1002/qj.2297](https://doi.org/10.1002/qj.2297)
- Cowtan K, Hausfather Z, Hawkins E, Jacobs P, Mann ME, Miller SK, Steinman BA, Stolpe MB, Way RG (2015) Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophys Res Lett* 42:6526–6534. doi:[10.1002/2015GL064888](https://doi.org/10.1002/2015GL064888)
- England MH, McGregor S, Spence P, Meehl GA, Timmermann A, Cai W, Santos A et al (2014) Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nat Clim Change* 4(3):222–227
- England MH, Kajtar JB, Maher N (2015) Robust warming projections despite the recent hiatus. *Nat Clim change* 5:394–396. doi:[10.1038/nclimate2575](https://doi.org/10.1038/nclimate2575)
- Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M (2013) Evaluation of climate models. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Folland CK, Parker DE, Colman AW, Washington R (1999) Large scale modes of ocean surface temperature since the late nineteenth century. In: Navarra A (ed) *Beyond El Nino: decadal and interdecadal climate variability*. Springer, Verlag, pp 73–102
- Fyfe JC, Gillett NP (2014) Recent observed and simulated warming. *Nat Clim Change* 4:150–151
- Fyfe JC, Gillett FW, Zwiers NP (2013) Overestimated global warming over the past 20 years. *Nat Clim Change* 3:767–769
- Fyfe JC, Meehl GA, England MH, Mann ME, Santer BD, Flato GM, Hawkins Ed, Gillett NP, Xie SP, Kosaka Y, Swart NC (2016) Making sense of the early-2000s warming slowdown. *Nat Clim Change* 6:224–228. doi:[10.1038/nclimate2938](https://doi.org/10.1038/nclimate2938)
- Goddard L (2014) Heat hide and seek. *Nat Clim Change* 4:158–161
- Han W, Meehl GA, Hu A, Alexander MA, Yamagata T, Yuan D, Ishii M, Pegion P, Zheng J, Hamlington BD, Quan X, Leben RR (2013) Intensification of decadal and multi-decadal sea level variability in the western tropical Pacific during recent decades. *Clim Dyn*. doi:[10.1007/s00382-013-1951-1](https://doi.org/10.1007/s00382-013-1951-1)
- Hawkins E, Sutton R (2015) Connecting climate model projections of global temperature change with the real world. *Bull Am Meteorol Soc*. doi:[10.1175/BAMS-D-14-00154.1](https://doi.org/10.1175/BAMS-D-14-00154.1)
- Hawkins E, Edwards T, McNeall D (2014) Pause for thought. *Nat Clim Change* 4:154–156
- Henley BJ, Gergis J, Karoly DJ, Power SB, Kennedy J, Folland CK (2015) A tripole index for the Interdecadal Pacific Oscillation. *Clim Dyn*. doi:[10.1007/s00382-015-2525-1](https://doi.org/10.1007/s00382-015-2525-1)
- Huber M, Knutti R (2014) Natural variability, radiative forcing and climate response in the recent hiatus reconciled. *Nat Geosci* 7(9):651–656
- IPCC (2013) Summary for policymakers. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Jones GS, Stott PA, Christidis N (2013) Attribution of observed historical near surface temperature variations to anthropogenic and natural causes using CMIP5 simulations. *J Geophys Res Atmos*. doi:[10.1002/jgrd.50239](https://doi.org/10.1002/jgrd.50239)
- Karl TR, Arguez A, Huang B, Lawrimore JH, McMahon JR, Menne MJ, Peterson TC, Vose RS, Zhang H-M (2015) Possible artifacts of data biases in the recent global surface warming hiatus. *Science* 26:1469–1472. doi:[10.1126/science.aaa5632](https://doi.org/10.1126/science.aaa5632)
- Kirtman B, Power SB, Adedoyin JA, Boer GJ, Bojariu R, Camilloni I, Doblas-Reyes FJ, Fiore AM, Kimoto M, Meehl GA, Prather M, Sarr A, Schär C, Sutton R, van Oldenborgh GJ, Vecchi G, Wang HJ (2013) Near-term climate change: projections and predictability. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Kociuba G, Power SB (2015) Inability of CMIP5 models to simulate recent strengthening of the walker circulation: implications for projections. *J Clim* 28:20–35. doi:[10.1175/JCLI-D-13-00752.1](https://doi.org/10.1175/JCLI-D-13-00752.1)
- Kosaka Y, Xie S-P (2013) Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* 501(7467):403–407

- Lewandowsky S, Risbey JA, Oreskes N (2015) On the definition and identifiability of the alleged “hiatus” in global warming. *Sci Rep*. doi:[10.1038/srep16784](https://doi.org/10.1038/srep16784) (**Article number: 16784**)
- Lewandowsky S, Risbey JA, Oreskes N (2016) The “Pause” in global warming: turning a routine fluctuation into a problem for science. *Bull Am Meteorol Soc*. doi:[10.1175/BAMS-D-14-00106.1](https://doi.org/10.1175/BAMS-D-14-00106.1)
- Marotzke J, Forster PM (2015) Forcing, feedback and internal variability in global temperature trends. *Nature* 517(7536):565–570
- Meehl GA, Arblaster JM (2012) Relating the strength of the tropospheric biennial oscillation (TBO) to the phase of the Interdecadal Pacific Oscillation (IPO). *Geophys Res Lett* 39:L20716. doi:[10.1029/2012GL053386](https://doi.org/10.1029/2012GL053386)
- Meehl et al (2011) Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat Clim Change* 1:360–364
- Meehl GA, Hu A, Arblaster JM, Fasullo J, Trenberth KE (2013) Externally forced and internally generated decadal climate variability associated with the Interdecadal Pacific Oscillation. *J Clim* 26:7298–7310. doi:[10.1175/JCLI-D-12-00548.1](https://doi.org/10.1175/JCLI-D-12-00548.1)
- Meehl GA, Teng H, Arblaster JM (2014) Climate model simulations of the observed early-2000s hiatus of global warming. *Nat Clim Change* 4:898–902. doi:[10.1038/nclimate2357](https://doi.org/10.1038/nclimate2357)
- Moss RH et al (2010) The next generation of scenarios for climate change research and assessment. *Nature* 463:747–756. doi:[10.1038/nature08823](https://doi.org/10.1038/nature08823)
- Newman M, Compo GP, Alexander MA (2003) ENSO-forced variability of the Pacific decadal oscillation. *J Clim* 16:3853–3857
- Power SB, Smith IN (2007) Weakening of the walker circulation and apparent dominance of El Niño both reach record levels, but has ENSO really changed? *Geophys Res Lett*. doi:[10.129/2007/GL30854](https://doi.org/10.129/2007/GL30854)
- Power S, Casey T, Folland C, Colman A, Mehta V (1999) Interdecadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:234–319
- Power S, Haylock M, Colman R, Wang X (2005) The predictability of interdecadal changes in ENSO activity and ENSO teleconnections. *J Clim* 19:4755–4771
- Power SB, Delage F, Colman R, Moise A (2012) Consensus of twenty-first century rainfall projections in climate models more widespread than previously thought. *J Clim* 25:3792–3809. doi:[10.1175/JCLI-D-11-00354.1](https://doi.org/10.1175/JCLI-D-11-00354.1)
- Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res*. doi:[10.1029/2002JD002670](https://doi.org/10.1029/2002JD002670)
- Risbey JS, Lewandowsky S, Langlais C, Monselesan DP, O’Kane TJ, Oreskes N (2014) Well-estimated global surface warming in climate projections selected for ENSO phase. *Nat Clim Change* 4(9):835–840
- Salinger MJ et al (2001) Interdecadal Pacific Oscillation and South Pacific climate. *Int J Climatol*. doi:[10.1002/joc.691](https://doi.org/10.1002/joc.691)
- Santer BD et al (2014) Volcanic contribution to decadal changes in tropospheric temperature. *Nat Geosci* 7:185–189
- Schmidt GA, Shindell DT, Tsigaridis K (2014) Reconciling warming trends. *Nat Geosci* 7:158–160. doi:[10.1038/ngeo2105](https://doi.org/10.1038/ngeo2105)
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93(4):485–498
- Trenberth KE, Fasullo JT (2013) An apparent hiatus in global warming? *Earth’s Future* 1:19–32. doi:[10.1002/2013EF000165](https://doi.org/10.1002/2013EF000165)
- van Vuuren DP et al (2011) The representative concentration pathways: an overview. *Clim Change* 109:5–31. doi:[10.1007/s10584-011-0148-z](https://doi.org/10.1007/s10584-011-0148-z)
- Vose RS, Arndt D, Banzon VF, Easterling DR, Gleason B, Huang B, Kearns E, Lawrimore JH, Menne MJ, Peterson TC, Reynolds RW, Smith TM, Williams CN Jr, Wuertz DL (2012) NOAA’s merged land-ocean surface temperature analysis. *Bull Am Meteorol Soc* 93:1677–1685. doi:[10.1175/BAMS-D-11-00241.1](https://doi.org/10.1175/BAMS-D-11-00241.1)
- Watanabe M, Shiogama H, Tatebe H, Hayashi M, Ishii M, Kimoto M (2014) Contribution of natural decadal variability to global warming acceleration and hiatus. *Nat Clim Change* 4(10):893–897
- Wehner MF, Easterling DR (2015) The global warming hiatus’s irrelevance. *Science* 350:1482–1483